JOURNAL OF VIROLOGY, Feb. 1997, p. 996–1003 Vol. 71, No. 2
0022-538X/97/$04.00+0

The Internal Open Reading Frame within the Nucleocapsid Gene of Mouse Hepatitis Virus Encodes a Structural Protein That Is Not Essential for Viral Replication

FRANÇOISE FISCHER,[1] DING PENG,[1]† SUSAN T. HINGLEY,[3]‡ SUSAN R. WEISS,[3] AND PAUL S. MASTERS[1,2]*

Department of Biomedical Sciences, State University of New York at Albany, Albany, New York 12237[1]; Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany, New York 12201[2]; and Department of Microbiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104[3]

Received 19 August 1996/Accepted 25 October 1996

The coronavirus mouse hepatitis virus (MHV) contains a large open reading frame embedded entirely within the 5′ half of its nucleocapsid (N) gene. This internal gene (designated I) is in the +1 reading frame with respect to the N gene, and it encodes a mostly hydrophobic 23-kDa polypeptide. We have found that this protein is expressed in MHV-infected cells and that it is a previously unrecognized structural protein of the virion. To analyze the potential biological importance of the I gene, we disrupted its expression by site-directed mutagenesis using targeted RNA recombination. The start codon for I was replaced by a threonine codon, and a stop codon was introduced at a short interval downstream. Both alterations created silent changes in the N reading frame. In vitro translation studies showed that these mutations completely abolished synthesis of I protein, and immunological analysis of infected cell lysates confirmed this conclusion. The MHV I mutant was viable and grew to high titer. However, the I mutant had a reduced plaque size in comparison with its isogenic wild-type counterpart, suggesting that expression of I confers some minor growth advantage to the virus. The engineered mutations were stable during the course of experimental infection in mice, and the I mutant showed no significant differences from wild type in its ability to replicate in the brains or livers of infected animals. These results demonstrate that I protein is not essential for the replication of MHV either in tissue culture or in its natural host.

Coronaviruses are a family of enveloped, single-stranded, positive-sense RNA viruses that employ a variety of unusual, in some cases unique, strategies for gene expression at both the transcriptional and translational levels (19, 34). They also hold the distinction of having the largest genomes among all of the RNA viruses. Mouse hepatitis virus (MHV), the murine member of the coronavirus family, provides a useful laboratory model for viral pathogenesis, particularly of virus-induced neurological diseases and hepatitis (2, 36). A major advantage of MHV as a subject of study is that both virus and host are tractable to genetic as well as biochemical analysis. The virions of MHV are constituted of a nucleocapsid (N) protein that binds the 31-kb genomic RNA to form a helically symmetric nucleocapsid, which is surrounded by a membrane envelope containing two or three major glycoproteins: the spike (S), the membrane (M), and, in some strains, the hemagglutinin-esterase (HE) proteins. A recently recognized minor structural protein, the small membrane (sM) protein, also resides in the MHV membrane (40).

The 5′ two-thirds of the MHV genome is devoted to a single gene encoding a huge polyprotein, the final products of which are thought to carry out genome transcription and replication. The genes for the structural proteins are clustered in the 3′ third of the genome, with the N gene as the most distal. Interspersed among the structural protein genes are a number of smaller open reading frames (ORFs). A detailed molecular understanding of MHV and other coronaviruses will require, in part, an evaluation of the significance of these ORFs.

One of the incompletely characterized ORFs is found embedded entirely within the 5′ half of the N gene. This internal (I) gene is in the +1 reading frame relative to N and encodes a largely hydrophobic polypeptide of 203 to 220 amino acids. It appears in 10 of the 11 N genes of different MHV strains that have been sequenced to date (3, 9, 16, 17, 24, 35) as well as in the rat coronavirus sialodacryoadenitis virus (16) and in bovine coronavirus (BCV), where it was first described (20). One line of reasoning prevalent in RNA virology proposes that such ORFs must have a purpose that has caused them to be maintained throughout evolution. Otherwise, an accumulation of stop codons in frames other than the primary reading frame would be expected owing to the relatively high error rate of RNA-dependent RNA polymerases. Indeed, examination of all potential coding regions in frames other than the primary reading frames in the remainder of the MHV genome reveals no ORFs even approaching the size of the I gene.

In several RNA viruses, examples of mRNAs with overlapping reading frames have been found (29). The earliest demonstrations that such messages are functionally bicistronic were made for influenza B virus (33), Sendai virus (7), and reovirus (4, 11). In most, but not all, of these cases, the translation of two products from the same mRNA can be explained by a leaky-scanning mechanism in which ribosomes occasionally bypass the first AUG because it falls in a suboptimal context and initiate at a downstream AUG, which is in a more favorable context (14, 15). This arrangement pertains for the N and I genes of MHV strain A59 (MHV-A59) as well as BCV. For BCV, it has been shown that the I protein is expressed in

* Corresponding author. Mailing address: David Axelrod Institute, Wadsworth Center, NYSDOH, New Scotland Ave., P.O. Box 22002, Albany, NY 12201-2002. Phone: (518) 474-1283. Fax: (518) 473-1326.

† Present address: Memorial Sloan Kettering Cancer Center, New York, NY 10021.

‡ Present address: Department of Microbiology and Immunology, Philadelphia College of Osteopathic Medicine, Philadelphia, PA 19131.

infected cells and is membrane associated (32). Moreover, hyperimmune serum from BCV-infected calves recognizes I protein, indicating that I is expressed during infection in the natural host. However, it is unknown whether the I protein has any significant function in coronavirus infection. This point is of interest by itself, and it becomes crucial in consideration of prospective genetic analyses of the N gene.

MATERIALS AND METHODS

Virus and cells. Wild-type, mutant, and recombinant virus stocks of MHV-A59 were propagated in mouse 17 clone 1 (17Cl1) cells. Plaque titer determinations and plaque purifications were carried out in mouse L2 cells. Spinner cultures of L2 cells, used for electroporation-mediated RNA transfection, were maintained as previously described (23).

Plasmid constructs. A donor plasmid for the expression of a synthetic MHV-A59 defective interfering (DI) RNA under the control of the T7 RNA polymerase promoter was prepared by extension of a prior DI RNA-encoding plasmid, pB36 (23). This was accomplished by transfer of half of the M gene from plasmid pCK1 (25) in a three-way ligation made up of the (blunted) *Acc*65I-*Apa*I fragment of pCK1 and the larger (blunted) *Bam*HI-*Bsi*WI and *Bsi*WI-*Apa*I fragments of pB36 (see Fig. 2). Although pCK1 is a cDNA clone from the genome of an MHV mutant (25), the fragment used comprises only wild-type sequence. The resulting construct, pP17, contains the first 466 nucleotides (nt) of the 5′ end of the MHV genome directly fused to the last 1,973 nt of the 3′ end of the MHV genome (plus a polyadenylate tail). The latter includes half of the M gene, the M-N intergenic sequence, the N gene, and the 3′ untranslated region (UTR). The fusion between the partial gene 1 and the partial M gene is in frame and potentially encodes a chimeric polypeptide of 182 amino acids.

Mutagenesis of the I gene within pP17, to disrupt the start codon and to create a new internal stop codon, was carried out by splicing overlap extension (SOE)-PCR (10). Nucleotides 2 (T) and 11 (C) in the I gene were mutated to, respectively, C and A. Two intermediate SOE-PCR fragments were amplified from pCK1 template by using mutagenic oligonucleotides PM199 (5′ TAACGGAAT CCTAAAGAAGACCACT 3′; mutations underlined) and PM200 (5′ CTTTA GGATTCCGTTACCAGCGCGG 3′; mutations underlined) paired with N-gene primer DP4 (5′ AGAGCGGGTATTGGTGTC 3′) and M-gene primer PM149 (5′ GATTACCATACACTAACA 3′) (see Fig. 2). These products were then purified, combined, and amplified with the outside primers, DP4 and PM149, and the final SOE-PCR product was restricted with *Age*I and *Apa*I and inserted into pP17 via a three-way ligation with the larger *Bsi*WI-*Apa*I and the smaller *Bsi*WI-*Age*I fragments of pP17. The resulting plasmid, pFF11, is identical to pP17 with the exception of the two specified point mutations.

Transcription vectors for the production of mRNAs for in vitro translation were derived from plasmid pA50, a transcription vector for the MHV N gene under the control of the T7 RNA polymerase promoter (22). The mRNA encoded by pA50 contains the start codon of the N gene (first AUG) in reading frame 0 and an intermediate start codon (second AUG) and the start codon of the I gene (third AUG) in the +1 reading frame. A construct deleting the region encompassing the first and second AUG, pFF14, was made by insertion of the (blunted) *Bsr*FI-*Apa*I fragment of pA50 back into the large (blunted) *Kpn*I-*Apa*I fragment of pA50 (see Fig. 5). The mRNA encoded by pFF14 thus has the I-gene start codon as its first AUG. A second plasmid, pFF16, identical to pFF14 except for the two point mutations that disrupt the I gene, was made in the same way, using pFF11 as the source of the inserted (blunted) *Bsr*FI-*Apa*I fragment (see Fig. 5). In both pFF14 and pFF16, the expected (blunted) *Kpn*I-(blunted) *Bsr*FI junction (G/CCGGT) was found to lack one base (the underlined C); however, this preceded the I coding region and was not of consequence for the purpose of the constructs. A third plasmid obtained from pA50 was pFF41, in which the second AUG was mutated to ACG such that the start codon of the I gene became the second AUG in the encoded mRNA. This mutant was generated by production of an 80-bp PCR fragment from pA50 template with a positive-sense T7 promoter primer (5′ AATACGACTCACTATAG 3′) and a negative-sense mutagenic primer FF9 (5′ CTTCTGCCACCGGCGTTTTCTTGCCCAGGA 3′; mutation underlined). The resulting PCR product was restricted with *Kpn*I and *Bsr*FI and used in a three-way ligation with the *Bsr*FI-*Apa*I and the large *Kpn*I-*Apa*I fragments of pA50 (see Fig. 5).

Recombinant DNA manipulations were carried out by standard methods (28). The compositions of all constructs were verified by restriction analysis; all newly generated junctions and PCR-created segments were confirmed by DNA sequencing by the method of Sanger et al. (30), using modified T7 DNA polymerase (Sequenase; U.S. Biochemical).

Targeted recombination. The two point mutations in the I gene were transduced into MHV by the targeted RNA recombination method described previously (13, 23), using MHV mutant Alb1 (23) as the recipient virus and in vitro-synthesized transcripts from *Nsi*I-truncated pFF11 as donor RNA. In parallel, an isogenic wild-type recombinant virus was generated from Alb1 by the same procedure, using pP17-derived transcripts as donor RNA. Alb1 infection and DI RNA transfection of mouse L2 spinner cells, followed by plating onto 17C11 cell monolayers, were performed as described in detail previously (23). Recombinants were directly selected, without prior heat treatment, as viruses forming large (i.e., wild-type-size) plaques at the nonpermissive temperature (39°C) and were plaque purified at this temperature. Final confirmation of recombinant construction was established by sequencing of the relevant regions of purified viral genomic RNA, using primers complementary to nt 142 to 159 of the N gene (for the I-gene mutations) and to nt 469 to 486 of the N gene (for the Alb1 mutations).

Virus purification, RNA purification, synthesis, and sequencing. Virus was purified by polyethylene glycol precipitation followed by two cycles of equilibrium centrifugation on preformed gradients of 0 to 50% potassium tartrate (osmotically counterbalanced with 30 to 0% glycerol) containing 50 mM Tris-maleate (pH 6.5)–1 mM EDTA. Gradients were centrifuged for 16 to 24 h at 111,000 × *g* in a Beckman SW41 rotor at 4°C. Following each cycle of equilibrium centrifugation, viral bands were collected from gradients, diluted with 50 mM Tris-maleate (pH 6.5)–100 mM NaCl–1 mM EDTA, and pelleted onto glycerol cushions by centrifugation for 2 h at 151,000 × *g* in a Beckman SW41 rotor at 4°C. Isolation of viral genomic RNA was done exactly as described previously (13). For the isolation of total cytoplasmic RNA from virus-infected 17Cl1 cell monolayers, a Nonidet P-40–gentle lysis procedure was used (12). Direct RNA sequencing was carried out by a modification of a dideoxy termination protocol (5, 25). Synthetic capped DI RNAs and mRNAs were produced with phage T7 RNA polymerase in standard reactions (22) or with a transcription kit (Ambion) as instructed by the manufacturer, and RNA products were purified by using RNaid (Bio 101, Inc.).

In vitro translation. Synthetic capped transcripts from *Hin*dIII-linearized pFF14, pFF16, pA50, and pFF41 were translated in a micrococcal nuclease-treated rabbit reticulocyte lysate (Amersham). Approximately 1 μg of each mRNA was used to program 15 μl of reticulocyte lysate containing 16 U of RNasin and 10 μCi of [^{35}S]methionine, in a final volume of 20 μl. Incubations were carried out for 90 min at 30°C. In vitro-synthesized proteins (4 μl of each sample) were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 12.5% polyacrylamide gel (18). Gels were fixed in 50% methanol–10% acetic acid and impregnated with 1 M sodium salicylate prior to drying and fluorography at −70°C (1).

Antiserum preparation and Western blot analysis. For the generation of antipeptide antisera, two peptides were synthesized on an Applied Biosystems 413A synthesizer. The first, MESSRRPLGLTKPSVDQIIKC, corresponded to the amino-terminal 20 amino acids of the predicted I-gene product followed by an additional cysteine residue to allow one-site chemical coupling; the second, PSSKRERSFSLQKDKEC, corresponded to residues 52 to 68 of I protein. Coupling to ovalbumin was carried out with sulfosuccinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (Pierce Chemical) by a previously described procedure (37). Rabbits were immunized subcutaneously at multiple sites on the back with a total of 1 mg of conjugated peptide emulsified with Freund's complete adjuvant, and three boosts were performed similarly at 1-month intervals with 1 mg of conjugated peptide in Freund's incomplete adjuvant. Rabbits were bled 2 weeks after each inoculation.

Confluent 20-cm^2 monolayers of 17Cl1 cells were mock infected or infected with Alb110 or Alb111 at a multiplicity of 10 PFU/cell. At 17 h postinfection, monolayers were washed twice with phosphate-buffered saline and lysed with 500 μl of 1× immunoprecipitation buffer (10 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM EDTA, 0.25% Nonidet P-40, 0.2 mg of phenylmethylsulfonyl fluoride per ml, 0.01% sodium azide) for 15 min at 4°C. Lysates were clarified by centrifugation, and 15 μl of each was separated by SDS-PAGE on 14% polyacrylamide, transferred to a polyvinylidine difluoride membrane (Millipore), and probed with antisera specific for I protein or N protein. Bound antibodies were visualized with an enhanced chemiluminescence detection system (Amersham).

Analysis of infectivity in mice. MHV-free C57BL/6 mice 4 to 6 weeks of age (Jackson Laboratory) were inoculated intracerebrally with 4×10^3 PFU of Alb110 (I^-) or Alb111 (I^+) passage 3 virus. At 2, 4, 6, and 8 days postinfection (dpi), duplicate mice infected with each virus were sacrificed, and infectious titers were determined for brain and liver homogenates on L2 cells at 37°C (8). For some samples from Alb110-infected mice, homogenates from brain and liver were then used to infect monolayers of 17Cl1 cells, and total cytoplasmic RNA isolated from second-passage infected cells was sequenced in the region of the engineered mutations.

RESULTS

Generation of an I-gene mutant of MHV by targeted RNA recombination. To study the contribution of the I gene to the biology of MHV-A59, we disrupted this gene by introducing two point mutations into its ORF. The first point mutation abolished the I-gene start codon by converting the AUG into ACG, while the second point mutation, 9 nt downstream, created a stop codon by converting UCA into UAA (Fig. 1). These two mutations were designed to be silent within the N ORF (the first point mutation interconverts two synonymous asparagine codons; the second interconverts two leucine codons), and therefore they should not interfere with N-gene expression.

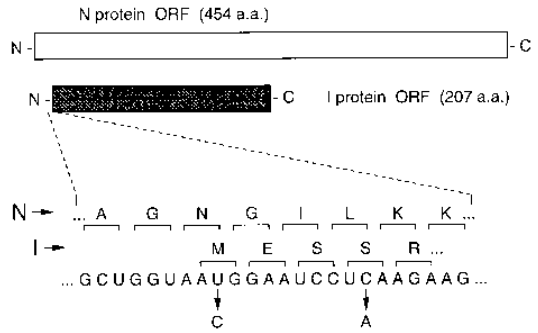

FIG. 1. Strategy for specific disruption of the MHV-A59 I gene. The positions of the overlapping ORFs for the N and I proteins are indicated at the top. The I ORF is in the +1 reading frame relative to the N ORF. Shown below are the two nucleotide changes introduced (at nt 2 and 11 in the I gene or nt 66 and 75 in the N gene), which disrupt I but are silent in N. a.a., amino acids.

An I-gene mutant and an isogenic wild-type recombinant control were constructed by means of a targeted RNA recombination technique described previously (13, 23), using an MHV N-gene mutant, Alb1 (23), as the recipient virus. Alb1 is a temperature-sensitive and thermolabile mutant that forms tiny plaques at the nonpermissive temperature (39°C) but is indistinguishable from the wild type at the permissive temperature (33°C). The N gene of Alb1 contains two closely spaced point mutations (within an interval of 10 nt); only the second of these is responsible for the phenotype of the mutant (23). The Alb1 mutations are some 300 nt downstream of the locus of the two mutations that we sought to transduce into the I gene. To produce donor DI RNAs for targeted RNA recombination, two vectors, pP17 and pFF11, were constructed (Fig. 2). All vectors that we have previously used to introduce mutations into the N gene and 3′ UTR of MHV contain the entire N gene fused to sequence corresponding to the 5′ end of the MHV genome (13, 23, 24, 26). However, since the I-gene mutations are very close to the 5′ end of the N gene, the minimal wild-type DI RNA vector, pB36, was expanded for the purposes of this study. The insertion of half of the M gene and the M-N intergenic sequence into pB36 produced wild-type vector pP17, from which was derived pFF11, containing the I-gene mutations (Fig. 2). This extension provided more flanking sequence in the region 5′ to the I-gene mutations, with the intent of increasing the probability of a productive crossover event (Fig. 3).

Synthetic donor RNAs transcribed from pFF11 and pP17 were transfected independently into Alb1-infected cells, and recombinants among the progeny were isolated as viruses able to form large plaques at the nonpermissive temperature. Initially, one recombinant constructed from each donor RNA was chosen for further analysis: an I-gene mutant virus, Alb110 (from pFF11), and its isogenic wild-type counterpart, Alb111 (from pP17). RNA sequencing analysis of the relevant portions of purified genomic RNA from Alb110 and Alb111 demonstrated the presence of the two expected I-gene mutations in Alb110 and the corresponding wild-type sequence in Alb111 (Fig. 4A). Similarly, it was found that the two initial mutations from Alb1 had been replaced by wild-type residues in both Alb110 and Alb111 (Fig. 4B). These results strongly suggest that each recombinant was formed by a single homologous recombination event (Fig. 3) that occurred downstream of the gene 1–M-gene discontinuity in the donor RNA and upstream

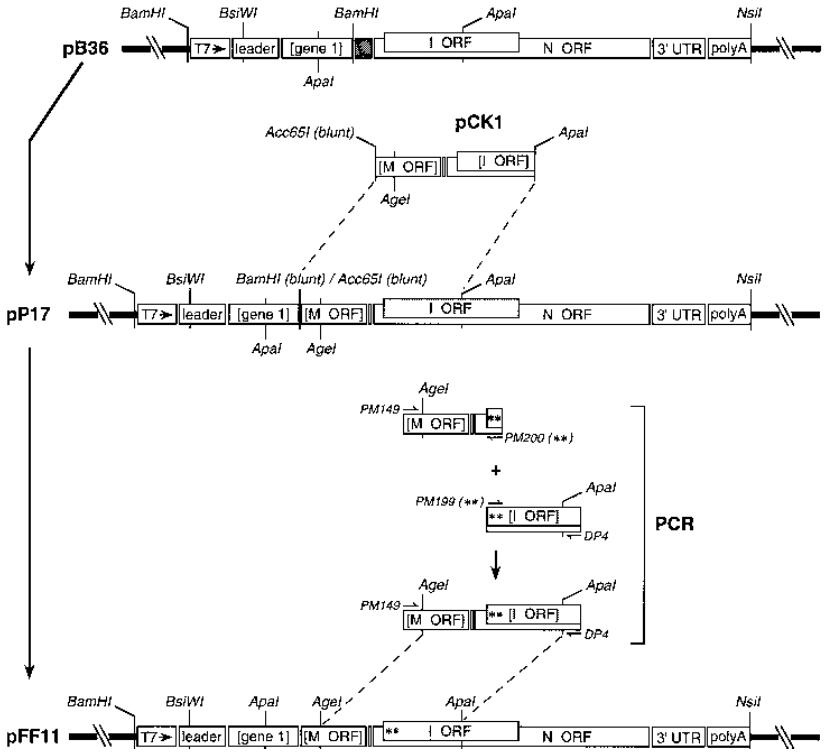

FIG. 2. Construction of transcription vectors used for synthesis of MHV DI RNAs containing the entire N gene as well as upstream flanking sequence from the M gene. Plasmid pP17 contains the wild-type sequence; pFF11 includes the two point mutations that disrupt the I gene (represented by asterisks). Restriction sites relevant to the construction of the plasmids, as detailed in Materials and Methods, are indicated. Regions shown in brackets are gene fragments rather than entire genes.

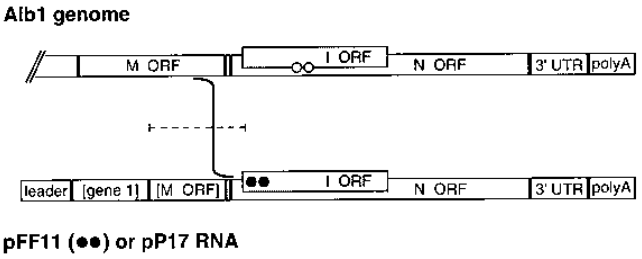

FIG. 3. Targeted recombination between the Alb1 genome and synthetic DI RNA to construct a mutant containing two point mutations in the I gene. Open circles represent two point mutations in Alb1, which change both the N and I ORFs (23). Closed circles represent the two engineered point mutations (present in pFF11 but absent in pP17) that disrupt the I ORF but are silent in the N ORF. The dotted line indicates the region in which a single crossover event could generate the selected mutants, Alb110 and Alb126 to Alb128 (with pFF11 RNA as the donor RNA), or a wild-type isogenic recombinant, Alb111 (with pP17 RNA as the donor RNA).

of the two I-gene mutations (for Alb110) or upstream of the Alb1 mutations (for Alb111).

Tissue culture analysis of the I-gene mutant. Since it was possible to obtain I-gene mutants and since these arose at frequencies comparable to those of wild-type control recombinants, it was clearly established that the disruption of the I gene is not lethal to MHV with respect to growth in tissue culture. Indeed, passage 2 stocks of Alb110 and Alb111 were found to have infectious titers of 2.0×10^8 and 2.3×10^8 PFU/ml, respectively. These results indicated that there were no gross differences in the replication characteristics of these two recombinants. However, examination of the plaque sizes of Alb110 and Alb111 at 33, 37, and 39°C revealed that the plaque size of the I-gene mutant was consistently smaller than that of its wild-type counterpart. This result suggested that I-gene expression does make a subtle contribution to MHV replication. Alternatively, it was possible that Alb110 had acquired a mutation other than those that we had engineered.

To examine whether Alb110 had actually been created by a nonhomologous recombination event occurring in the vicinity of the gene 1–M-gene discontinuity in the donor RNA, the region of the *Acc*65I site in the M gene of the I mutant was analyzed by sequencing of genomic RNA. This portion of the Alb110 M gene was found to be identical to the wild-type sequence (data not shown). This finding ruled out the possibility that a nonhomologous crossover event generated a small deletion or insertion in the M gene, thereby producing the observed Alb110 phenotype.

To address the possibility that the Alb110 phenotype was due to some spontaneously arising mutation other than the two point mutations that we introduced into the I gene, we isolated three additional independent I mutant viruses. The new recombinants, designated Alb126, Alb127, and Alb 128, were obtained from separate transfections and were verified by the same criteria used for Alb110. These three isolates were analyzed in parallel with Alb110, Alb111, and wild-type MHV-A59 in plaque assays performed at 33, 37, and 39°C. As shown in Table 1, all of the I-gene mutants produced smaller plaques on L2 cells than the isogenic wild-type control virus, Alb111, which had plaques identical in size to those of wild-type MHV-A59. Moreover, this effect was independent of temperature (Table 1), and it was also seen for plaques formed on 17Cl1 and Sac^- cells. The appearance of the same phenotype in four independently isolated I mutants argues strongly that the observed phenotype is indeed associated with the two introduced point mutations and not with other mutations localized elsewhere in the genome. This result indicates that the I gene, although not essential in tissue culture, contributes to the plaque-forming efficiency of MHV.

Infectivity of the I-gene mutant in mice. To further investigate the potential contribution of the I gene to the tissue tropism and infectivity of MHV-A59, in vivo experiments were performed. Mice inoculated intracerebrally with Alb110 or Alb111 were sacrificed at 2-day intervals for 8 days, and viral infectious titers were determined for homogenates prepared from brain and liver. The results presented in Table 2 show that each virus could be reisolated from both the brains and the livers of infected mice at almost all times postinfection examined, and the titers of the various reisolates were not substantially different. The average liver titers for Alb111 showed a trend toward slightly higher values (ca. 0.5 $\log_{10}$) than those for

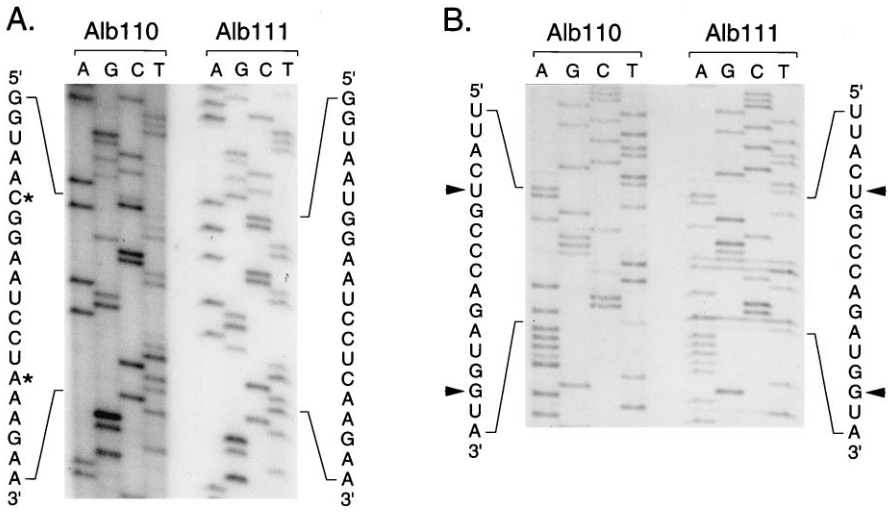

FIG. 4. Nucleotide sequence of genomic RNA isolated from purified virus of I gene mutant Alb110 or isogenic wild-type control Alb111. The indicated segments of sequence are given as positive-sense RNA. (A) Sequence of the region where point mutations were introduced into the I gene. The two point mutations in Alb110 are denoted by asterisks. (B) Sequence of the region of the N gene containing the Alb1 mutations. Arrowheads denote positions of the two original mutations in Alb1 that have been restored to wild-type residues in Alb110 and Alb111.

TABLE 1. Plaque sizes of recombinants at different temperatures[a]

Virus	Plaque size (mm)		
	33°C	37°C	39°C
Alb110 (I^-)	0.5	1.8	1.4
Alb126 (I^-)	0.7	1.8	1.6
Alb127 (I^-)	0.5	1.6	1.5
Alb128 (I^-)	0.5	1.6	1.4
Alb111 (I^+)	1.2	2.8	2.3
Wild type (I^+)	1.0	2.6	2.3

[a] Plaque diameters were measured at 48 h after infection of L2 cell monolayers. Each value is the average from 4 to 12 plaques.

Alb110, but it is not certain that this is significant. Clearly, the I mutant and its wild-type counterpart showed no marked difference in ability to replicate in the brain or in the liver or in ability to spread to the liver after intracerebral inoculation.

To confirm the stability of the two mutations of Alb110 during the course of this experimental infection, various brain and liver reisolates of Alb110 were passaged twice in tissue culture, and infected cellular RNA was subsequently sequenced. A total of five samples were examined: two each for brain at 2 dpi, two each for brain at 8 dpi, and one for liver at 4 dpi. The two expected I-disrupting mutations were found in all samples (data not shown), indicating that the infectious titers given for Alb110 in Table 2 are truly those of the mutant and do not represent a revertant population selected for during infection in the host. Moreover, the plaque size phenotype of Alb110 was retained for virus reisolated after passage in mice. Altogether, these data demonstrate that, at least under the experimental conditions used, the tissue tropism and in vivo infectivity of MHV-A59 are not dependent on the expression of the I gene.

In vitro translation of the I gene in different contexts. To ascertain whether the two point mutations introduced into the I gene did, in fact, disrupt the expression of this gene, in vitro translation experiments were carried out. For this purpose, we constructed a transcription vector, pFF14, in which the I gene is the first ORF in the encoded mRNA (Fig. 5). This allowed unequivocal identification of the I-protein product without ambiguities introduced by products related to the N protein being cotranslated from the same message. In addition, this optimized translation of I protein, thereby providing the most stringent test of the efficacy of the I-gene-disrupting mutations, which were incorporated into an otherwise identical plasmid, pFF16 (Fig. 5). Capped, runoff mRNAs were synthesized from

TABLE 2. In vivo replication of Alb110 (I ORF mutant) and Alb111 (isogenic wild-type control)[a]

Organ	Day post-infection	Titer (PFU/g)	
		Alb110	Alb111
Brain	2	5.6×10^4, 5.0×10^6	2.1×10^7, 1.2×10^6
	4	2.3×10^7, 6.5×10^7	8.9×10^7, 9.3×10^6
	6	3.7×10^5, 1.0×10^6	7.2×10^5, 5.1×10^5
	8	1.3×10^5, 4.6×10^4	3.8×10^4, 1.9×10^4
Liver	2	1.6×10^4, 8.4×10^3	9.4×10^3, 9.1×10^4
	4	6.7×10^4, 7.7×10^4	2.3×10^6, 2.8×10^4
	6	5.5×10^3, 6.1×10^2	8.2×10^4, 3.2×10^3
	8	$<5 \times 10^2$, $<5 \times 10^2$	$<5 \times 10^2$, $<5 \times 10^2$

[a] Mice were inoculated intracerebrally with 4×10^3 PFU of either virus. Duplicate animals were sacrificed at the indicated times, and viral titers in brain and liver were determined on mouse L2 cells at 37°C.

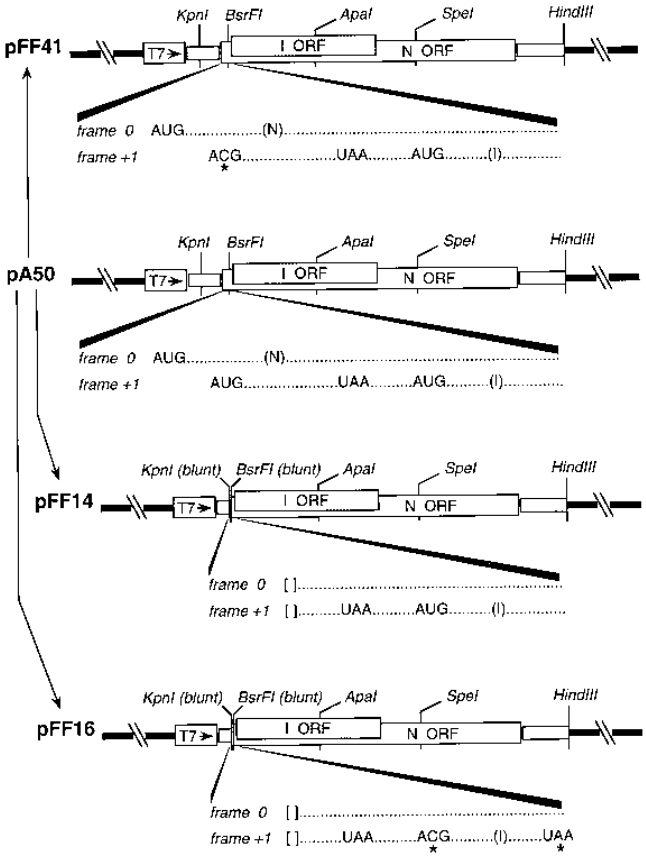

FIG. 5. Construction of transcription vectors used for synthesis of mRNAs for in vitro translation. Plasmid pA50, described previously (22), acts as the template for a runoff transcript containing the entire N (and I) gene in addition to most of the 5′ leader and 3′ UTR. Restriction sites relevant to the construction of the other plasmids, each derived from pA50 as detailed in Materials and Methods, are indicated. Beneath each vector is shown the region of the encoded mRNA containing the N initiation codon (in reading frame 0) and the initiation codons for I and a short ORF situated between the starts of N and I (in the +1 reading frame). Asterisks denote point mutations; brackets denote deletions. In pFF14, the first two start codons have been deleted, placing the I start codon closest to the 5′ end of the transcript. Plasmid pFF16 has the same deletion as pFF14 but also contains the two point mutations previously used to disrupt the I gene. Plasmid pFF41 is identical to pA50 except that the second start codon, between those of N and I, has been mutated.

*Hin*dIII-truncated pFF14 and pFF16 and were used to program a rabbit reticulocyte lysate. SDS-PAGE analysis of the translation product from pFF14-derived mRNA showed primarily a 23-kDa polypeptide (Fig. 6, lane 4), which agrees well with the calculated size of the I protein (22.6 kDa). In many in vitro translation experiments, we have observed this band as a doublet. It is not clear whether this is due to proteolytic processing or posttranslational modification. Significantly, absolutely no I protein was detected by translation of pFF16-derived mRNA (Fig. 6, lane 5). These results demonstrate that the two introduced mutations totally abolish I-gene expression.

Minor amounts of two smaller species (20 and 14 kDa) were observed in the translation products of mRNAs generated from both pFF14 and pFF16. These were also seen if the transcription vectors were truncated with *Spe*I instead of *Hin*dIII (Fig. 5). The precise origin of these polypeptides has not been determined. The 14-kDa product possibly originated from a downstream AUG within the I ORF (at nt 213). However, the 20-kDa product cannot be accounted for by initiation at any downstream AUG within the I gene or within the N

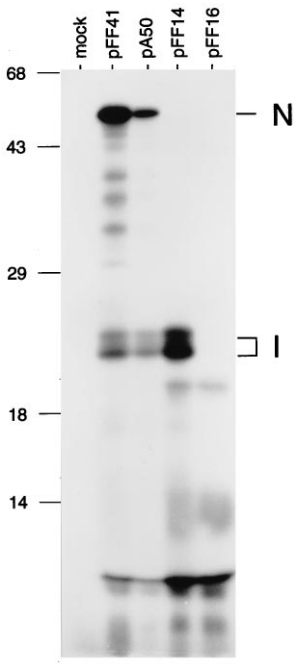

FIG. 6. In vitro translation of mRNAs to examine the effects of different start codon contexts on I-protein expression. Synthetic transcripts were translated in a rabbit reticulocyte lysate, as described in Materials and Methods, and [^{35}S]methionine-labeled protein products were analyzed by SDS-PAGE followed by fluorography. Numbers on the left indicate the molecular masses (in kilodaltons) of marker proteins.

gene, and there are no ORFs of significant size in the +2 reading frame. We cannot rule out a combination of non-AUG initiation and abortive termination or frameshifting as potential sources of these minor products.

The same set of experiments was used to address an ancillary issue regarding I-protein expression. With the exception of MHV-A59 and MHV-1, in all MHV strains and in BCV, the I ORF begins at the second AUG in the +1 reading frame relative to N. In MHV-A59 (and MHV-1), an intermediate AUG, also in the +1 reading frame, potentially initiates a short (octapeptide) ORF preceding the I-gene start codon. To assess whether the intermediate AUG has an impact on the expression of I protein, we constructed a plasmid, pFF41, in which the second AUG has been disrupted such that the start codon of the I ORF becomes the second AUG in the encoded mRNA (Fig. 5). Translation of pFF41-derived mRNA was compared with that of mRNA from its parent plasmid, pA50 (Fig. 5), which contains the three ORFs in their unaltered wild-type context. Both the wild-type and the mutant mRNAs programmed synthesis of both N and I proteins (Fig. 6, lanes 2 and 3). Furthermore, the ratio of translated I protein to N protein was not increased by disruption of the intermediate ORF. Therefore, it is unlikely that the short ORF affects expression of I protein during MHV-A59 infection.

In vivo expression of I protein. Antiserum specific for I protein was obtained by immunization of a rabbit with a peptide corresponding to the amino-terminal 20 amino acids of the predicted I-gene product. In Western blot analysis, this antiserum recognized a 23-kDa polypeptide in reticulocyte lysate that had been programmed with pFF14-derived mRNA (Fig. 7A), and this band had the same mobility as [^{35}S]methionine-labeled, in vitro-translated I protein in the same lysate (Fig. 7B). Also, the 23-kDa band was not recognized by serum obtained from the same rabbit prior to immunization, nor was it recognized by the antipeptide antiserum in reticulocyte lysate that had not been programmed with I mRNA (Fig. 7A and B). This established the specificity of this immunological reagent for I protein.

Western blot analysis of lysates prepared from infected cells revealed that the antipeptide antiserum, but not preimmune serum, recognized a protein in Alb111-infected cells having the same mobility as in vitro-translated I protein. However, the same protein was absent from Alb110-infected cells or mock-infected cells (Fig. 7C). Identical results were obtained when a second antipeptide antibody that had been raised against an internal segment of the I protein was used (data not shown). In addition, equal amounts of N protein were detected in the lysates from Alb110- and Alb111-infected cells, indicating that the infections by each virus had proceeded to corresponding extents (Fig. 7D). Thus, I protein was clearly expressed during wild-type MHV infection, and this expression was abrogated in the engineered I-gene mutant.

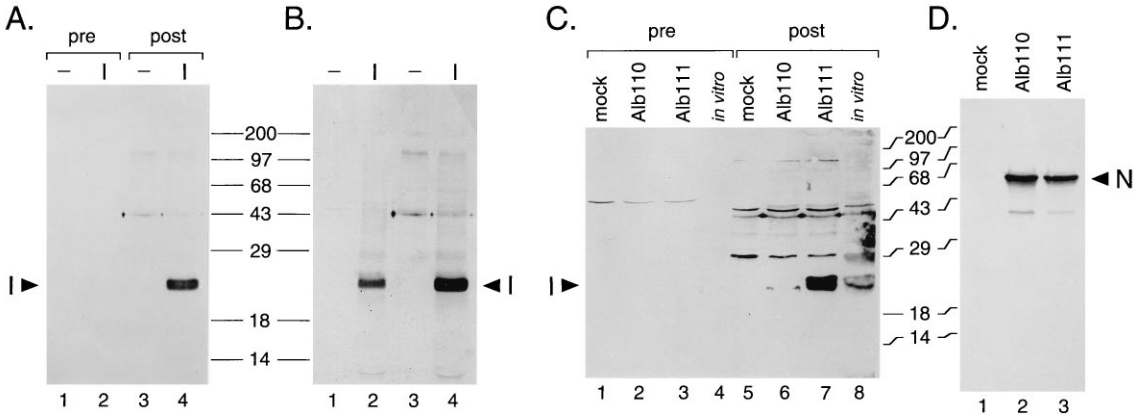

FIG. 7. Analysis of I-protein expression in vivo. (A) Samples of [^{35}S]methionine-labeled reticulocyte lysates programmed with H_2O (lanes 1 and 3) or pFF14-derived mRNA (lanes 2 and 4) were separated by SDS-PAGE and probed in a Western blot with preimmune serum (lanes 1 and 2) or with I-protein antipeptide antiserum (lanes 3 and 4). The blot was visualized by detection of chemiluminescence (30-s exposure). (B) Autoradiogram (48-h exposure) of the same blot as in panel A following decay of the chemiluminescent signal. (C) Western blot of lysates from mock-, Alb110-, or Alb111-infected cells or in vitro-translated I protein probed with preimmune serum (lanes 1 to 4) or with I-protein antipeptide antiserum (lanes 5 to 8). (D) Western blot of lysates from mock-, Alb110-, or Alb111-infected cells probed with MHV N-protein-specific antisera. Molecular masses (in kilodaltons) of marker proteins are indicated.

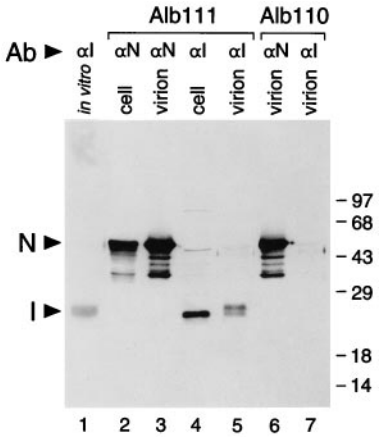

FIG. 8. Incorporation of I protein into virions. A lysate from Alb111-infected cells (lanes 2 and 4) or virions of Alb111 (lanes 3 and 5) or Alb110 (lanes 6 and 7) purified by two cycles of equilibrium centrifugation on potassium tartrate-glycerol gradients were analyzed by Western blotting and visualized by detection of chemiluminescence. Also included was a control of in vitro-translated I protein (lane 1). Samples were probed with I-protein antipeptide antiserum (αI; lanes 1, 4, 5, and 7) or with N-protein-specific antisera (αN; lanes 2, 3, and 6). Molecular masses (in kilodaltons) of marker proteins are indicated. Ab, antibody.

I protein is an MHV structural protein. Purified virions of Alb110 and Alb111 were isolated by two cycles of equilibrium centrifugation on potassium tartrate-glycerol gradients as described in Materials and Methods. When the I-protein-specific antipeptide antiserum was used in Western blot analysis of purified Alb111, it was found, unexpectedly, that I protein is a constituent of MHV virions (Fig. 8, lane 5). Consistent with this observation, no such protein was detected in virions of Alb110 (Fig. 8, lane 7). Since the quantities of N protein in the cell lysate and virion samples in this Western blot were roughly equivalent (Fig. 8, lanes 2 and 3), it can be seen that I protein appeared in assembled virions in a proportion comparable to the fraction of intracellular N protein that was incorporated into virions. It could be argued that the hydrophobic I protein may have been present in cell membrane debris that copurified with virions. However, we think that this possibility is very unlikely since the procedure used to isolate MHV virions yields a highly pure preparation, which, in previous radiolabeling experiments, has been clearly devoid of cellular proteins (13, 26). Therefore, we conclude that I protein is a previously unrecognized structural protein of MHV.

DISCUSSION

We demonstrate in this report that the 23-kDa polypeptide potentially encoded by the internal ORF in the +1 reading frame within the N gene is, indeed, expressed in MHV-infected cells. In addition, the I protein has been found to be a constituent of isolated virions. Further, we describe the generation of an MHV mutant in which the I ORF has been disrupted by targeted RNA recombination. The options available for construction of a null mutant in this gene were severely constrained by the complete overlap of the I gene with the gene for an essential structural protein, N. While the changes introduced into the I ORF are silent in the N ORF, in vitro translation results supported the assumption that these mutations completely abolished expression of I protein (Fig. 6), and this conclusion was confirmed by Western blot analysis of infected cells and virions (Fig. 7 and 8). The viability of the I mutant, independently isolated four times, demonstrated that the mutations we have introduced do not block growth of MHV in tissue culture. Furthermore, following intracerebral inoculation of mice, mutant virus could be recovered from both brain and liver tissue at titers comparable to those of the wild type in parallel infections. Altogether these results establish that the I protein is not essential for replication of MHV either in tissue culture or during experimental infection of its target species. The I gene thus joins genes 2a, HE, 4, and 5a as MHV genes whose expression has been shown to not be required for replication in mouse cells and, in some cases, in the animal host (21, 31, 38, 39).

It is noteworthy that disruption of the I gene brought about a modest but significant effect on virus growth in tissue culture. The four independent isolates of the I mutant all had a markedly smaller plaque size in multiple cell lines, and this phenotype was found to be independent of temperature. In this respect, there is an interesting parallel with the C gene of measles virus (MV). Many paramyxoviruses contain a gene, designated C, that is located wholly within the first half of the gene for the viral phosphoprotein, and, like I, C occurs in the +1 reading frame and is thought to be translated via a leaky-scanning mechanism. It has recently been shown, through the construction of an engineered mutant, that disruption of the C gene of MV does not affect the ability of virus to grow to high titer in tissue culture, although the phenotype of the MV C^- mutant has not yet been examined in infected animals (27). Curiously, the MV C^- mutant also produces smaller plaques than its isogenic wild-type counterpart in tissue culture. In the case of MHV, we believe that the small-plaque characteristic reflects the changes made in the I gene and was not caused by unintended changes in the N gene. The synonymous substitutions made in the N gene do not create rare codons, nor do they affect the context of the N initiation codon, since they are some 70 nt downstream. There recently has been reported a novel effect on translation of the Sindbis virus capsid protein caused by sequence downstream of the initiation codon, which has the potential to form an RNA secondary structure (6). If there were a functionally analogous mechanism in MHV, this would raise the possibility that the I mutations alter an RNA secondary structure that somehow contributes to N-gene expression. However, this possibility is ruled out by the observation that in infected cells, there is no difference in N-protein synthesis between the I mutant and its isogenic wild-type counterpart, as assessed by either radiolabeling (data not shown) or Western blotting of N protein (Fig. 7D).

A secondary conclusion of this study is that a potential source of ambiguity in genetic analysis of the MHV N gene can be eliminated. Since the I gene is contained within the 5′ half of the N gene, this could have complicated the interpretation of the effect of mutations in this portion of the N gene. Thus, our results are also relevant for future structure and function analyses of the amino-terminal end of the N protein, as they demonstrate that mutations can be introduced into the 5′ half of the N gene with, at most, minor phenotypic effects attributable to the I gene.

The finding that the I protein is a structural component of the MHV virion was surprising, in light of the amount of prior attention that has been given to characterization of this virus. We hypothesize that the reason I protein has been previously overlooked in isotopically labeled virions is that it may be masked by the multiple isoforms of the very abundant M protein, which migrate to the same apparent molecular weight range on SDS-PAGE. In purified virions, I protein migrated as two bands, only the lower one of which was detected in soluble extracts from infected cells (Fig. 8, lanes 4 and 5). Two poten-

tial explanations for this observation can be proposed. First, I protein may undergo a rapid partial proteolysis in the cell lysate, whereas incorporation into virions stabilizes it in its original higher-molecular-weight form. Alternatively, I protein may undergo posttranslational modification to a slower-migrating species concomitant with its incorporation into virions. Further study will be required to resolve this matter.

At this point, the identification of the biochemical function of the I protein remains speculative. We have not found growth conditions that either exaggerate or diminish the plaque size phenotype. Many viruses encode auxiliary gene products that counteract specific components of the host antiviral response induced by interferon. However, the I mutant and its isogenic wild-type counterpart are equally sensitive to the effects of low doses of interferon (data not shown). Although it is nonessential, I protein may augment some rate-limiting process in viral replication such that its absence results in smaller plaques. Alternatively, the plaque size phenotype may not reflect the rate of viral replication but could be related to impairment of virus spreading by cell-cell fusion, a mechanism mediated by the S protein. Thus, conceivably the I protein plays an enhancing role in the intracellular processing or transport of S protein. The tissue culture phenotype of the I mutant may also suggest that there remains to be found a more subtle differential aspect to its infectivity in the host, perhaps via a different mode of inoculation or in a different strain of mouse.

ACKNOWLEDGMENTS

We are grateful to James Dias and James Seeger of the Wadsworth Center Peptide Synthesis Facility for the synthesis of peptides and for valuable advice. We thank Tim Moran and Matthew Shudt of the Molecular Genetics Core Facility of the Wadsworth Center for the synthesis of oligonucleotides. We are grateful to Xiurong Wang for expert technical assistance.

This work was supported in part by Public Health Service grants AI 31622 (P.S.M.), NS 30606 (S.R.W.), and NS 21954 (S.R.W.) from the National Institutes of Health. F.F. received support from the Fondation Mérieux.

REFERENCES

1. **Chamberlain, J. P.** 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. Anal. Biochem. **98:**132–135.
2. **Compton, S. R., S. W. Barthold, and A. L. Smith.** 1993. The cellular and molecular pathogenesis of coronaviruses. Lab. Anim. Sci. **43:**15–28.
3. **Décimo, D., H. Philippe, M. Hadchouel, M. Tardieu, and M. Meunier-Rotival.** 1993. The gene encoding the nucleocapsid protein: sequence analysis in murine hepatitis virus type 3 and evolution in *Coronaviridae*. Arch. Virol. **130:**279–288.
4. **Ernst, H., and A. J. Shatkin.** 1985. Reovirus hemagglutinin mRNA codes for two polypeptides in overlapping reading frames. Proc. Natl. Acad. Sci. USA **82:**48–52.
5. **Fichot, O., and M. Girard.** 1990. An improved method for sequencing of RNA templates. Nucleic Acids Res. **18:**6162.
6. **Frolov, I., and S. Schlesinger.** 1994. Translation of Sindbis virus mRNA: effects of sequences downstream of the initiating codon. J. Virol. **68:**8111–8117.
7. **Giorgi, C., B. M. Blumberg, and D. Kolakofsky.** 1983. Sendai virus contains overlapping genes expressed from a single mRNA. Cell **35:**829–836.
8. **Hingley, S. T., J. L. Gombold, E. Lavi, and S. R. Weiss.** 1994. MHV-A59 fusion mutants are attenuated and display altered hepatotropism. Virology **200:**1–10.
9. **Homberger, F. R.** 1995. Sequence analysis of the nucleoprotein genes of three enterotropic strains of murine coronavirus. Arch. Virol. **140:**571–579.
10. **Horton, R. M., and L. R. Pease.** 1991. Recombination and mutagenesis of DNA sequences using PCR. *In* M. J. McPherson (ed.), Directed mutagenesis, a practical approach. IRL Press, New York, N.Y.
11. **Jacobs, B. L., and C. E. Samuel.** 1985. Biosynthesis of reovirus-specified polypeptides: the reovirus s1 mRNA encodes two primary translation products. Virology **143:**63–74.
12. **Kingsman, S. M., and C. E. Samuel.** 1980. Mechanism of interferon action. Interferon-mediated inhibition of simian virus-40 early RNA accumulation. Virology **101:**458–465.
13. **Koetzner, C. A., M. M. Parker, C. S. Ricard, L. S. Sturman, and P. S. Masters.** 1992. Repair and mutagenesis of the genome of a deletion mutant of the coronavirus mouse hepatitis virus by targeted RNA recombination. J. Virol. **66:**1841–1848.
14. **Kozak, M.** 1986. Bifunctional messenger RNAs in eukaryotes. Cell **47:**481–483.
15. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108:**229–241.
16. **Kunita, S., M. Mori, and E. Terada.** 1993. Sequence analysis of the nucleocapsid protein gene of rat coronavirus SDAV-681. Virology **193:**520–523.
17. **Kunita, S., E. Terada, K. Goto, and N. Kagiyama.** 1992. Sequence analysis and molecular detection of mouse hepatitis virus using the polymerase chain reaction. Lab. Anim. Sci. **42:**593–598.
18. **Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–688.
19. **Lai, M. M. C.** 1990. Coronavirus: organization, replication and expression of genome. Annu. Rev. Microbiol. **44:**303–333.
20. **Lapps, W., B. G. Hogue, and D. A. Brian.** 1987. Sequence analysis of the bovine coronavirus nucleocapsid and matrix protein genes. Virology **157:**47–57.
21. **Luytjes, W., P. J. Bredenbeek, A. F. H. Noten, M. C. Horzinek, and W. J. M. Spaan.** 1988. Sequence of mouse hepatitis virus A59 mRNA2: indications for RNA recombination between coronaviruses and influenza C virus. Virology **166:**415–422.
22. **Masters, P. S.** 1992. Localization of an RNA-binding domain in the nucleocapsid protein of the coronavirus mouse hepatitis virus. Arch. Virol. **125:**141–160.
23. **Masters, P. S., C. A. Koetzner, C. A. Kerr, and Y. Heo.** 1994. Optimization of targeted RNA recombination and mapping of a novel nucleocapsid gene mutation in the coronavirus mouse hepatitis virus. J. Virol. **68:**328–337.
24. **Parker, M. M., and P. S. Masters.** 1990. Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. Virology **179:**463–468.
25. **Peng, D., C. A. Koetzner, and P. S. Masters.** 1995. Analysis of second-site revertants of a murine coronavirus nucleocapsid protein deletion mutant and construction of nucleocapsid protein mutants by targeted RNA recombination. J. Virol. **69:**3449–3457.
26. **Peng, D., C. A. Koetzner, T. McMahon, Y. Zhu, and P. S. Masters.** 1995. Construction of murine coronavirus mutants containing interspecies chimeric nucleocapsid proteins. J. Virol. **69:**5475–5484.
27. **Radecke, F., and M. A. Billeter.** 1996. The nonstructural C protein is not essential for multiplication of Edmonston B measles virus in cultured cells. Virology **217:**418–421.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
29. **Samuel, C. E.** 1989. Polycistronic animal virus mRNAs. Prog. Nucleic Acid Res. Mol. Biol. **37:**127–153.
30. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
31. **Schwarz, B., E. Routledge, and S. G. Siddell.** 1990. Murine nonstructural protein ns2 is not essential for virus replication in transformed cells. J. Virol. **64:**4784–4791.
32. **Senanayake, S. D., M. A. Hofmann, J. L. Maki, and D. A. Brian.** 1992. The nucleocapsid protein gene of bovine coronavirus is bicistronic. J. Virol. **66:**5277–5283.
33. **Shaw, M. W., P. W. Choppin, and R. A. Lamb.** 1983. A previously unrecognized influenza B virus glycoprotein from a bicistronic mRNA that also encodes the viral neuraminidase. Proc. Natl. Acad. Sci. USA **80:**4879–4883.
34. **Siddell, S. G.** 1995. The Coronaviridae: an introduction, p. 1–10. *In* S. G. Siddell (ed.), The Coronaviridae. Plenum Press, New York.
35. **Skinner, M. A., and S. G. Siddell.** 1983. Coronavirus JHM: nucleotide sequence of the mRNA that encodes nucleocapsid protein. Nucleic Acids Res. **11:**5045–5054.
36. **Sturman, L. S., and K. V. Holmes.** 1983. The molecular biology of coronaviruses. Adv. Virus Res. **28:**35–111.
37. **Weiner, R. S., T. T. Andersen, and J. A. Dias.** 1990. Topographic analysis of the alpha-subunit of human follicle-stimulating hormone using site-specific antipeptide antisera. Endocrinology **127:**573–579.
38. **Weiss, S. R., P. W. Zoltick, and J. L. Leibowitz.** 1993. The ns 4 gene of mouse hepatitis virus (MHV), strain A59 contains two ORFs and thus differs from ns 4 of the JHM and S strains. Arch. Virol. **129:**301–309.
39. **Yokomori, K., and M. M. C. Lai.** 1991. Mouse hepatitis virus S RNA sequence reveals that nonstructural proteins ns4 and ns5a are not essential for murine coronavirus replication. J. Virol. **65:**5605–5608.
40. **Yu, X., W. Bi, S. R. Weiss, and J. L. Leibowitz.** 1994. Mouse hepatitis virus gene 5b protein is a new virion envelope protein. Virology **202:**1018–1023.